Table 1: Alpha1a SNPs

SNP Numbr (FINAL)	SNP location on1, B=Exi /ar4, D=Spl E=SplVar2	SNP Sequence (nucleotide context)	Nucleotide Change A=reference sequence a=SNP sequence (A/a)	Repeats reference n=	Repeats2 (range) n= (when applicable)	Type of nucleotide Δ	Location In Gene	Amino acid change?	Amino acid2 change?	Conservative amino acid change?	Type of2 amino subst	Δ Restriction Site? yes/no
	Exon 1											
1	A3465	CAATCAGATTTG	G/T			transversion	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
2	A3472	TTTGTCATATTT	C/T			transversion	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
3	A3501	TATTTAAAACAC	A/C			transversion	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
4	A3719	TACAAATGACT	A/G			transition	Exon1, 5' reg	no Δ aa	no Δ aa			yes
5	A3740 I/D	AC repeat	(AC)n	n=7	n=5-8	insert/del	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
6	A3754 I/D	GC repeat	(GC)n	n=7	n=6-12	insert/del	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
7	A3776 I/D	AC repeat	(AC)n	n=11	n=8-15	insert/del	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
	A3727-76	...	(AC/GC/AC)n	n=25	n=20-30	insert/del	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
8	A3828	GAGGAAAACATT	A/G			transition	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
9	A3928	TCAGCGAGAGC	C/A			transversion	Exon1, 5' reg	no Δ aa	no Δ aa			yes
10	A4048	TTTTAAAAAATG	A/T			transversion	Exon1, 5' reg	no Δ aa	no Δ aa			yes
11	A4094	GACATCAGTGGTGG	A/G			transition	Exon1, 5' reg	no Δ aa	no Δ aa			yes
12	A4109	GGGGTGAGTCAGC	A/G			transition	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
13	A4215	GGTCCGCATCC	G/A			transition	Exon1, 5' reg	no Δ aa	no Δ aa			yes
14	A4315	CCTCCCCGCCGT	C/G			transversion	Exon1, 5' reg	no Δ aa	no Δ aa			yes
15	A4385	ATACCGGCCCT	G/A			transition	Exon1, 5' reg	no Δ aa	no Δ aa			yes
16	A4634	CCTGCCGAGTC	C/T			transition	Exon1, 5' reg	no Δ aa	no Δ aa			yes
17	A4777	CTGTCGCTTGC	G/A			transition	Exon1, 5' reg	no Δ aa	no Δ aa			yes
18	A4890	GAAAGGCGTCATG	G/A			transition	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
19	A4893	GGCGTCATGGAC	A/G			transition	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
20	A4956	TGGCTCTCAAGCC	T/C			transition	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
21	A4965	GCCAGCTTGGCT	C/T			transition	Exon1, 5' reg	no Δ aa	no Δ aa			yes
22	A5027	GAGAATGAGCG	A/G			transition	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
23	A5291	TTTTTTT	(T)n	n=7	n=7-9	insert/del	Exon1, 5' reg	no Δ aa	no Δ aa			no Δ
24	A5662	CTCAAATGTA	A/G			transition	Exon1, 5'UTR	no Δ aa	no Δ aa			no Δ
25	A5859	GGAATTGCAT	T/C			transition	Exon1, 5'UTR	no Δ aa	no Δ aa			no Δ
26	A5996	TTTAATGCCCTG	C/T			transition	Exon1, 5'UTR	no Δ aa	no Δ aa			yes
27	A6158	AGCCCGGGAGGTGC	G/C			transversion	Exon1, coding, NH2 terminus	no Δ aa	no Δ aa			no Δ
28	A6220	CTCTCCCGAAATG	G/C			transversion	Exon1, coding, TM4	S to A	S154A	non-conservative	polar/neutral	no Δ
29	A6685	CACTCTCCCTGGT	T/G			transversion	Exon1, coding, 2nd EC loop (end TM4)	R to K	R166K	conservative	basic/basic	no Δ
30	A6702	GGCTGGAGGCAGCC	G/A			transition	Exon1, coding, TM5	I to S	I200S	non-conservative	neutral/polar	no Δ
31	A6804	CCATCATCCT	T/G			transversion	Exon1, coding, 3rd intracell loop	G to E	G247[R]	non-conservative	neutral/acidic	no Δ
32	A6944	GGCAGCGGGATGGC	G/A			transition	Exon1, 3' intron	no Δ aa	no Δ aa			yes
33	A7141	CCTCGGCTTCTG	G/T			transversion	Exon1, 3' intron	no Δ aa	no Δ aa			yes
34	A7200	TTTCAAACAGCACA	A/C			transversion	Exon1, 3' intron	no Δ aa	no Δ aa			no Δ
35	A7273	TCCAGTCTCATTT	T/C			transition	Exon1, 3' intron	no Δ aa	no Δ aa			no Δ
	Exon 2 / Spl Var 1											
36	B267	AGGAATGG	A/G			transition	Exon2, 5' intron	no Δ aa	no Δ aa			no Δ
37	B322	CAAATGTGA	T/C			transition	Exon2, 5' intron	no Δ aa	no Δ aa			yes
38	B504	AATAGTATT	G/A			transition	Exon2, coding, TM7	V to I	V311I	conservative	neutral/neutral	yes
39	B612	TCTCCGCAGA	C/T			transition	Exon2, coding, prox COOH terminus	R to C	R347C	non-conservative	neutral disulfide/basic	yes
40	B776	TTCTCTTCCATG	T/G			transversion	Exon2, coding, mid COOH terminus	no aa Δ	no aa Δ			yes
41	B968	AGGAAGTCT	A/T			transversion	SplVar1, coding, distal COOH terminus	E to D	E465D	conservative	acidic/acidic	yes
42	B1016	GGTACCCACCC	C/A			transversion	SplVar1, 3' intron	no Δ aa	no Δ aa			yes
43	B1034	TTCTCGGAAGG	G/A			transition	SplVar1, 3' intron	no Δ aa	no Δ aa			yes
	Spl Var 4											
44	C563	ATGCAGAAGGGG	G/A			transition	SplVar4, 5' intron	no Δ aa	no Δ aa			no Δ
	Spl Var 3											
45	D563	ATTCAGTTC	G/A			transition	SplVar3, 5' intron	no Δ aa	no Δ aa			no Δ
46	D578	GTTTGGAGA	G/A			transition	SplVar3, 5' intron	no Δ aa	no Δ aa			yes
47	D611	ACAGGGAGC	G/C			transversion	SplVar3, 5' intron	no Δ aa	no Δ aa			yes
48	D754	TTGTGATCA	G/A			transition	SplVar3, 5' intron	no Δ aa	no Δ aa			yes
49	D901	GGGGAGCAT	G/A			transition	SplVar3, 5' intron	no Δ aa	no Δ aa			no Δ
	Spl Var 2	none										

Reference Restriction Site (site/none)	SNP2 Restriction Site (site/none)	SNP location2 A=Exon1, B=Exon2, C=SpVar4, D=SpVar3, E=SpVar2	f(-) ALL n=280 (incl Coriell)	f(-)2 ALL n=184 (no Coriell)	f(-)3 BLACK n=43 (no Coriell)	f(-)4 HISPANIC n=41 (no Coriell)	f(-)5 CAU n (no Coriell)
		Exon 1					
none	none	A3465	0.388	0.400	0.410	0.421	0.376
none	none	A3472	0.006	0.008	0.013	0.026	0.000
none	none	A3501	0.009	0.014	0.084	0.000	0.000
none	MaeIII	A3719	0.005	0.003	0.013	0.000	0.000
Cfol / MvnI / BssHII	none	A3740 I/D					
none	Cfol / MvnI / BssHII	A3764 I/D					
none	none	A3776					
see above	see above	A3727-76	0.020	0.025	0.000	0.013	0.041
none	none	A3828	0.008	0.009	0.031	0.000	0.006
none	none	A3928	0.002	0.003	0.000	0.000	0.000
Dral / Tru9I	Dral / Tru9I	A4048	0.002	0.003	0.000	0.000	0.000
none	none	A4094	0.002	0.003	0.013	0.013	0.000
none	none	A4109	0.103	0.093	0.141	0.183	0.052
none	none	A4215	0.058	0.056	0.083	0.038	0.051
BstYI	MvnI / Cfol	A4315	0.185	0.140	0.138	0.054	0.178
Cfr10I/MspI/HpaII/Sau96I/HaeIII	none	A4385	0.057	0.049	0.000	0.163	0.021
HinfI	Ddel / HinfI	A4634	0.130	0.118	0.265	0.175	0.051
MaeIII	none	A4777	0.002	0.003	0.000	0.000	0.005
Acyl	none	A4890	0.077	0.067	0.147	0.081	0.037
none	none	A4893	0.012	0.015	0.078	0.000	0.000
none	Bgll	A4956	0.002	0.003	0.000	0.000	0.000
Aful / Bgll	none	A4965	0.102	0.080	0.216	0.088	0.030
none	none	A5027	0.471	0.347	0.400	0.318	0.350
none	none	A5290	0.009	0.003	0.012	0.000	0.000
none	none	A5662	0.002	0.003	0.000	0.013	0.000
none	none	A5859	0.004	0.005	0.028	0.000	0.000
none	MspI/HpaII/ScrFI/NciI	A5996	0.002	0.003	0.000	0.000	0.005
/NciI/ScrFI(2nd)/SmaI/MspI/Hpe	BseAI/MroI/HpaII/MspI	A6158	0.007	0.005	0.000	0.013	0.005
Aval	EcoRII	A6220	0.002	0.005	0.000	0.000	0.005
EcoRII	Ital	A6665	0.005	0.005	0.000	0.000	0.010
Ital	EcoRII/BstXI	A6702	0.003	0.003	0.000	0.023	0.000
EcoRII/BstXI	none	A6804	0.055	0.043	0.000	0.136	0.043
none	BpuAl	A6944	0.003	0.003	0.000	0.000	0.005
none	none	A7141	0.003	0.003	0.000	0.000	0.008
none	none	A7200					
none	none	A7273					
		Exon 2 / Spl Var 1					
none	none	B267	0.012	0.016	0.070	0.000	0.000
none	HpyCH4IV	B322	0.083	0.063	0.081	0.038	0.045
none	Sspl	B504	0.002	0.000	0.000	0.000	0.000
AclI/BsmAI	Psll/Sfcl/HpyCH4V/BsmAI	B612	0.463	0.482	0.267	0.525	0.565
Earl	BstF5I	B776	0.002	0.003	0.012	0.000	0.000
none	Rsal	B988	0.030	0.021	0.023	0.025	0.020
BanI / NiaIV	Ddel / Hpy188I	B1016	0.002	0.003	0.012	0.000	0.000
Hpy188I		B1034	0.002	0.003	0.000	0.000	0.005
		Splice Variant 4					
none	none	C563	0.068	0.028	0.060	0.071	0.010
		Splice Variant 3					
none	none	D563	0.043	0.040	0.012	0.038	0.046
none	Taql	D578	0.006	0.006	0.028	0.000	0.000
none	Ital	D611	0.035	0.031	0.027	0.075	0.010
Sau3AI/Ndell/Bcll/Dpnl	none	D754	0.015	0.017	0.068	0.000	0.005
none	none	D901	0.006	0.006	0.028	0.000	0.000
		Splice Variant 2					
none	none	none					

Schwinn IDF-Table 2(A=exon1)

Range: Sequence 1 to 7902

DNA sequence listing (top strand / complementary bottom strand), numbered by position:

```
 10  TAAACCATGT / ATTTGGTACA
 20  TTTGGTATAA / AAACCATATT
 30  ATGTAATAAT / TACATTATTA
 40  AGAATGAGTT / TCTTACTCAA
 50  CATGATATAA / GTACTATATT
 60  TATAAGTGAA / ATATTCACTT
 70  AAGTAAGATA / TTCATTCTAT
 80  TATAATTATA / ATATTAATAT
 90  GGTACAGTAT / CCATGTCATA
100  GATTACACTC / CTAATGTGAG

110  TTTTATATAA / AAAATATATT
120  TGTATACAAT / ACATATGTTA
130  GTATAATATG / CATATTATAC
140  TGAAATATAT / ACTTTATATA
150  ATGTATGCAT / TACATACGTA
160  AATAAATACA / TTATTTATGT
170  TGTACATAAA / ACATGTATTT
180  ATATTTCCTG / TATAAAGGAC
190  GTATGTACAT / CATACATGTA
200  GTGTATATAT / CACATATATA

210  ATTTTTCCAT / TAAAAAGGTA
220  GTATATAATT / CATATATTAA
230  AATACACATT / TTATGTGTAA
240  AACAAAAAAT / TTGTTTTTTA
250  GAGAAAATAT / CTCTTTTATA
260  GACAAATATT / CTGTTTATAA
270  TAACAATAGC / ATTGTTATCG
280  TATTCTCCA / ATAAGAGGT
290  GTGGCAGGAT / CACCGTCCTA
300  TACAGATAAT / ATGTCTATTA

310  TTTATTTTC / AAAATAAAAG
320  ATACATTTCT / TATGTAAAGA
330  CTATTATGAG / GATAATACTC
340  TCCAAATGTG / AGGTTTACAC
350  CATAAATTTT / GTATTTAAAA
360  ATAATTAAAT / TATTAATTTA
370  ACAAGTCAAT / TGTTCAGTTA
380  AAGCATTTTA / TTCGTAAAAT
390  ATGTGCCTAT / TACACGGATA
400  ATGTGCCTAT / TACACGGATA

410  AACAGCCCAT / TTGTCGGGTA
420  CAGCAAGTGA / GTCGTTCACT
430  AGAGATGAGA / TCTCTACTCT
440  TAAAAATGTG / ATTTTTACAC
450  TCTGCACTTG / AGACGTGAAC
460  GTACACATAT / CATGTGTATA
470  TGACTTTCTT / ACTGAAAGAA
480  TCCTTCCAT / AGGAAGGTA
490  TTCCAGATGA / AAGGTCTACT
500  TATTGACAAA / ATAACTGTTT

510  GATATTTGGG / CTATAAACCC
520  GGGTTGATTA / CCCAACTAAT
530  ATATGCCTCA / TATACGGAGT
540  GCTTAGAGTT / CGAATCTCAA
550  GCTATGCGAG / CGATACGCTC
560  AGCAAGTCT / TCGTTCAGA
570  TATTGTTTA / ATAACAAAAT
580  AAGGAATTAT / TTCCTTAATA
590  TTGCTGATCA / AACGACTAGT
600  GCTATAGGCT / CGATATCCGA

610  GACTTGACCT / CTGAACTGGA
620  CATATTCCTA / GTATAAGGAT
630  CCATGATAGT / GGTACTATCA
640  CCCACAGTGT / GGGTGTCACA
650  AGGGATGGGG / TCCCTACCCC
660  CCAAGACTT / GGTTCTGAA
670  GGAGTATGTT / CCTCATACAA
680  ATTTTCTCAG / TAAAAGAGTC
690  ATTTTCTCAG / TAAAAGAGTC
700  GGCAATGAAG / CCGTTACTTC

710  ATTTGGAAAA / TAAACCTTTT
720  ATGATGGCAA / TACTACCGTT
730  AGACAGAATT / TCTGTCTTAA
740  ACTATCTCAT / TGATAGAGTA
750  CAAAGACATC / GTTTCTGTAG
760  CCTCAGCAGT / GGAGTCGTCA
770  ATCTGTGGGC / TAGACACCCG
780  TGTGCAGTGT / ACACGTCACA
790  TCTCAGTTCG / AGAGTCAAGC
800  ATAAAGGATT / TATTTCCTAA

810  AGAACACAAT / TCTTGTGTTA
820  GCTCTTTGG / CGAGAAACC
830  AGAGCTGTGA / TCTCGACACT
840  CTTGATACTG / GAACTATGAC
850  CATCAATACC / GTAGTTATGG
860  TTTCTGAGAA / AAAGACTCTT
870  TTGTTTTTTC / AACAAAAAAG
880  ATTTCTTGC / TAAAGAACG
890  CTCTTAACT / GAGAATTGA
900  TCTTAAGCCT / AGAATTCGGA

910  TAGGAGAATT / ATCCTCTTAA
920  AGTTGAAAAG / TCAACTTTTC
930  GGGGTAGATA / CCCCATCTAT
940  CTAACATTAA / GATTGTAATT
950  GTCTTCTACT / CAGAAGATGA
960  CAATCATAAA / GTTAGTATTT
970  CTGTCATTTG / GACAGTAAAC
980  TTCCAGAACA / AAGGTCTTGT
990  TTCCAGAACA / AAGGTCTTGT
1000 CAGCTCCTAA / GTCGAGGATT

1010 TTCCATTGTG / AAGGTAACAC
1020 TATTGTTTTC / ATAACAAAAG
1030 TAAGGGAATG / ATTCCCTTAC
1040 ATAGACAGAT / TATCTGTCTA
1050 TCTTTATTT / AGAAATAAAA
1060 TTTAAACCTC / AAATTTGGAG
1070 TAAGCCTACC / ATTCGGATGG
1080 ACACTTGCCG / TGTGAACGGC
1090 AGTTCCTCAC / TCAAGGAGTG
1100 TAGTCACTAA / ATCAGTGATT

1110 GAAAGTCCTG / CTTTCAGGAC
1120 CCAATCAATG / GGTTAGTTAC
1130 CATGGGTTTA / GTACCCAAAT
1140 TGTCCATTGC / ACAGGTAACG
1150 TCAGCTCTTC / AGTCGAGAAG
1160 TCCAATCAGA / AGGTTAGTCT
1170 CTCATTCCCC / GAGTAAGGGG
1180 CAGCATCCCT / GTCGTAGGGA
1190 GACACACCAC / CTGTGTGGTG
1200 TCTAAATGC / AGATTTACG
```

7902

AGATTTACG   CTGTGTGGTG   GTCGTAGGGA   GAGTAAGGGG   AGGTTAGTCT   AGTCGAGAAG   ACAGGTAACG   GTACCAAAT   GGTTAGTTAC   CTTTCAGGAC

```
1210 GGCTGCTGAT   1220 GGTTCACCTT   1230 CCTCACTTTT   1240 GTCTACAAAT   1250 CTCAATCCTG
     CCGACGACTA        CCAAGTGGAA        GGAGTGAAAA        CAGATGTTTA        GAGTTAGGAC

1260 CTGATTCCAC   1270 AAATCCTACA   1280 TCAAGCAATA   1290 TCATTTATG    1300 AGTCTTTCCA
     GACTAAGGTG        TTTAGGATGT        AGTTCGTTAT        AGTAAAATAC        TCAGAAAGGT

1310 CAACCACCCC   1320 TTCAGGGGAT   1330 TCTTCAATTT   1340 CTGTCACACC   1350 GGAAGTCTTC
     GTTGGTGGGG        AAGTCCCCTA        AGAAGTTAAA        GACAGTGTGG        CCTTCAGAAG

1360 AGAGTATCAC   1370 CCTCAGAGCC   1380 AGGCAAGAGG   1390 GACCCCGGCT   1400 AGGGTTTCAG
     TCTCATAGTG        GGAGTCTCGG        TCCGTTCTCC        CTGGGGCCGA        TCCCAAAGTC

1410 GCTTTAGAGA   1420 GTCCAGCTCT   1430 GACTCCTTT    1440 GGCCATAGGA   1450 CTAATGTGAT
     CGAAATCTCT        CAGGTCGAGA        CTGAGGAAAA        CCGGTATCCT        GATTACACTA

1460 ATGCCCACCT   1470 GGAGCCTGTG   1480 CCCTCCTTTC   1490 TAGACCATGC   1500 CCTGGGACTC
     TACGGGTGGA        CCTCGGACAC        GGGAGGAAAG        ATCTGGTACG        GGACCCTGAG

1510 AGAATCCCTT   1520 GCCCCAGATG   1530 GCCACACAAT   1540 CACTTTCAGG   1550 TCCATTCTCT
     TCTTAGGGAA        CGGGGTCTAC        CGGTGTGTTA        GTGAAAGTCC        AGGTAAGAGA

1560 CTGGGCAGAC   1570 AACATCACAA   1580 ATGTGTGTAC   1590 CCCAAGGCCT   1600 GAGGCCAAGA
     GACCCGTCTG        TTGTAGTGTT        TACACACATG        GGGTTCCGGA        CTCCGGTTCT

1610 AGGCAGCTTT   1620 CTGGCTGTAG   1630 GGGCTGAGGT   1640 GTTCACACAC   1650 ATTTGCATGG
     TCCGTCGAAA        GACCGACATC        CCCGACTCCA        CAAGTGTGTG        TAAACGTACC

1660 CCCCTCAAGA   1670 CAAAGAACAA   1680 GGGGGAAAGT   1690 GAGAAGAAAA   1700 GAAGCAGCCA
     GGGGAGTTCT        GTTTCTTGTT        CCCCCTTTCA        CTCTTCTTTT        CTTCGTCGGT

1710 GTGATCAGGG   1720 CCAGCTCTTG   1730 CAACTTAGCC   1740 ATGTTGGGTC   1750 ATTCTGATTA
     CACTAGTCCC        GGTCGAGAAC        GTTGAATCGG        TACAACCCAG        TAAGACTAAT

1760 AACCACTTAG   1770 CTCAAGTGTA   1780 GTGCTCAAGA   1790 CACTTAGCAC   1800 ATTCTCCAGC
     TTGGTGAATC        GAGTTCACAT        CACGAGTTCT        GTGAATCGTG        TAAGAGTCG

1810 TGAATTTACC   1820 AGTGTTCATG   1830 GACGACCTGG   1840 GTTAGAAATA   1850 TATTTCGCTA
     ACTTAAATGG        TCACAAGTAC        CTGCTGGACC        CAATCTTTAT        ATAAAGCGAT

1860 TAAAGTAGCA   1870 TACAAAATGA   1880 GCAGAAAGGG   1890 AGTTAATAAG   1900 ATTAATAATA
     ATTTCATCGT        ATGTTTTACT        CGTCTTTCCC        TCAATTATTC        TAATTATTAT

1910 GAGTTAGTGA   1920 ATATTATGAG   1930 CTGAGTTTT    1940 GAGAAACGTA   1950 ATTTCTTTCA
     CTCAATCACT        TATAATACTC        GACTCAAAAA        CTCTTTGCAT        TAAAGAAAGT

1960 CAACACTAAT   1970 AACAACCTTG   1980 TGGGGGTTCA   1990 TTGTCTCCCT   2000 TTAAAAATTA
     GTTGTGATTA        TTGTTGGAAC        ACCCCCAAGT        AACAGAGGGA        AATTTTTAAT

2010 GGAAACCAAG   2020 GCTTTGCCAT   2030 GGTCGCATAG   2040 GAGGGTCAGA   2050 ATAGCATCTT
     CCTTGGTTC         CGAAACGGTA        CCAGCGTATC        CTCCCAGTCT        TATCGTAGAA

2060 TATGACCCAG   2070 AGCATACTCG   2080 TCTCCACTCC   2090 ACCTACCCAT   2100 GTGTACAACT
     ATACTGGGTC        TCGTATGAGC        AGAGGTGAGG        TGGATGGGTA        CACATGTTGA

2110 CAGACACTTT   2120 CTGGGATGTC   2130 CACGTCAACT   2140 ATTCTTAAA    2150 GAGTAACCAA
     GTCTGTGAAA        GACCCTACAG        GTGCAGTTGA        TAAGAAATTT        CTCATTGGTT

2160 CAGATGGATA   2170 GTTTTCTGTT   2180 TGTGAATCAA   2190 TGGTAGGTGA   2200 CTGAAAAATT
     GTCTACCTAT        CAAAAGACAA        ACACTTAGTT        ACCATCCACT        GACTTTTTAA

2210 GGTTCTGAGA   2220 GGTCGTTTG    2230 CAAGGATTGA   2240 TGGTCACAGG   2250 CTGAGAAGCA
     CCAAGACTCT        CCAGCAAAAC        GTTCCTAACT        ACCAGTGTCC        GACTCTTCGT

2260 GATTTGAAAG   2270 ACCTACCTGC   2280 TAGCAGCATA   2290 AGAGCTGCTC   2300 TTCCTTATCT
     CTAAACTTTC        TGGATGGACG        ATCGTCGTAT        TCTCGACGAG        AAGGAATAGA

2310 TAGTATTAAC   2320 TAGTTAATTA   2330 TTGGAGGTGG   2340 GTGCAGGGGT   2350 GGATTATGTG
     ATCATAATTG        ATCAATTAAT        AACCTCCACC        CACGTCCCCA        CCTAATACAC

2360 TATTCTTAAT   2370 TGTTGTGAGAG  2380 TGGGAACTGG   2390 GAGTTACAAA   2400 GACTTTTGCA
     ATAAGAATTA        ACAACATCTC        ACCCTTGACC        CTCAATGTTT        CTGAAAACGT
```

2410 AGTTTCGACC / TCAAAGCTGG
2420 TTGCAGAGCT / AACGTCTCGA
2430 GAGCAATTTT / CTCGTTAAAA
2440 CAGTTGCTT / GTCAACGAAA
2450 GCTTGCTGAT / CGAACGACTA
2460 AGCACTGCT / TCGTGACGAA
2470 CCCTTATCTA / GGGAATAGAT
2480 CCATGGAACA / GGTACCTTGT
2490 CATCTTAATG / GTAGAATTAC
2500 AAGAATTTGC / TTCTTAAACG

2510 ATTCACAGCA / TAAGTGTCGT
2520 TCAGGTTAAT / AGTATGTTTT
2530 GAATACAAAA / CTTATGTTTT
2540 CAAAACAGTG / GTTTTGTCAC
2550 TATATCCCTC / ATATAGGGAG
2560 TGATGGATGG / ACTACCTACC
2570 GATTTCGGAA / CTAAAGCCTT
2580 GCACAGACAT / CGTGTCTGTA
2590 TATACACATA / ATATGTGTAT
2600 TTTGATGATA / AAACTACTAT

2610 AAGTACTAGA / TTCATGATCT
2620 AGTGCAGGGA / TCACGTCCCT
2630 ATTGAGGTCA / TAACTCCAGT
2640 AGCTTCCTCC / TCGAAGGAGG
2650 TAAGGGGACT / ATTCCCCTGA
2660 GAATCCCAGA / CTTAGGGTCT
2670 GAGAGCAGGT / CTCTCGTCCA
2680 GACTTAGTAA / CTGAATCATT
2690 TGAAGAGTGG / ACTCTTCACC
2700 AGCTGTCTGT / TCGACAGACA

2710 TCAACCAGGA / AGTTGGTCCT
2720 TGCTCCTCCT / ACGAGGAGGA
2730 ATGGCACGAA / TACCGTGCTT
2740 ATTCAGTTT / TAAGTCAAAA
2750 AAAAATATAT / TTTTTATATA
2760 TAAATTCAAA / ATTTAAGTTT
2770 TCAAATGTGT / AGTTTACACA
2780 TAGGTGTGAG / ATCCACACTC
2790 TCTTATCCC / AAGAATAGGG
2800 TACAGGTATG / ATGTCCATAC

2810 AGGCAGAGGT / TCCGTCTCCA
2820 GGAGGACTTT / CCTCCTGAAA
2830 GTATACAATA / CATATGTTAT
2840 GAGAAATAAA / CTCTTTATTT
2850 TACATATATT / ATGTATATAA
2860 AGGTCTTCCA / TCCAGAAGGT
2870 TGACATAGGA / ACTGTATCCT
2880 TTTACTGACC / AAATGACTGG
2890 CTCTCATGGG / GAGAGTACCC
2900 CATTCCTCTG / GTAAGGAGAC

2910 AGGCATTTTG / TCCGTAAAAC
2920 AGATTATTG / TCTAAATAAC
2930 CTATAAAAGA / GATATTTTCT
2940 GCCTCCCAAA / CGGAGGGTTT
2950 CATTATCTCA / GTAATAGAGT
2960 GTAATCATAT / CATTAGTATA
2970 GTATCATAT / CATTAGTATA
2980 TAATATGATT / ATTATACTAA
2990 TTGTTCACAG / AACAAGTGTC
3000 GAGAGAATTT / CTCTCTTAAA

3010 AAGTGCCACT / TTCACGGTGA
3020 GCTTAAAGTT / CGAATTTCAA
3030 ATCTCCTTGT / TAGAGGAACA
3040 TCCTAGGTTT / AGGATCCAAA
3050 AAGGAGACCT / TTCCTCTGGA
3060 AGTAAATAAG / TCATTTATTC
3070 AACATTCCAC / TTGTAAGGTG
3080 TTGTCTGCA / AACAGACGT
3090 TCAATAAAGA / AGTTATTTCT
3100 TGAAGATGA / ACTTTCTACT

3110 CTTAGGAGGT / GAATCCTCCA
3120 GGGAATTGGA / CCCTTAACCT
3130 GTGGGAAACA / CACCCTTTGT
3140 TTTTTCTATG / AAAAAGATAC
3150 TTCCCGATAT / AAGGGCTATA
3160 TCTGAAACAC / AGACTTTGTG
3170 ATGTGACTTT / TACACTGAAA
3180 ATTCAATCAC / TAAGTTAGTG
3190 AAGGTAAACA / TTCCATTTGT
3200 GATTATGTAA / CTAATACATT

3210 TTTACCAGAA / AAATGGTCTT
3220 AAAAAGTAAT / TTTTTCATTA
3230 AAGACTGGTG / TTCTGACCAC
3240 GTGCTAGGTT / CACGATCCAA
3250 TTCATACTCC / AAGTATGAGG
3260 AGCTATTAAT / TCGATAATTA
3270 GAATTAAAGA / CTTAATTTCT
3280 GAGTAACACT / CTCATTGTGA
3290 CCTGAAAGGA / GGACTTTCCT
3300 TACCATTTTC / ATGGTAAAAG

3310 TCAAGAAAC / AGTTCTTTTG
3320 TGGAAAGAT / ACCTTTTCTA
3330 TGTGTGGCAT / ACACACCGTA
3340 TTAAAAAATA / AATTTTTTAT
3350 CCAAACTCTG / GGTTTGAGAC
3360 TGGCCATAAT / ACCGGTATTA
3370 GCTCTTAAAA / CGGAGAATTT
3380 TTCATCTGTC / AAGTAGACAG
3390 TAAAGAAATT / ATTTCTTTAA
3400 AGAAGTGAAT / TCTTCACTTA

3410 CATATTAAAT / GTATAATTTA
3420 AAGGTTTAGA / TTCCAAATCT
3430 TATGTCCACT / ATACAGGTGA
3440 TTATCTTCCT / AATAGAAGGA
3450 GAAAATATAA / CTTTTATATT
3460 TTTCATTACA / AAAGTAATGT
3470 ATCAGATTTG / TAGTCTAAAC
3480 ATCAGATTTG / TAGTCTAAAC
3490 TCTGATTTTA / AGACTAAAAT
3500 CTTGCTATTT / GAACGATAAA

3510 AAAACACCTT / TTTTGTGGAA
3520 ATAAATTTACT / TATTAAATGA
3530 TGCATATTTA / ACGTATAAAT
3540 GAATTACAAT / CTTAATGTTA
3550 ATTCTTAATA / TAAGAATTAT
3560 TACTTCTTGA / ATGAAGAACT
3570 TCTTAACAAA / AGAATTGTTT
3580 ACCTAGGCCA / TGGATCCGGT
3590 AATGTTAATC / TTACAATTAG
3600 AAATCAAGCT / TTAGTTCGA

3610 GTTCAAAGTT CAAGTTTCAA
3620 ACTTTATAGC TGAAATATCG
3630 ACATTCCTAT TGTAAGGATA
3640 GAACACCA CTTGTGTGGT
3650 TACACACAGC ATGTGTGTCG
3660 AATATCTAGC TTATAGATCG
3670 AAGGGTGTCA TTCCCACAGT
3680 ATTTTTCGTT TAAAAAGCAA
3690 ATTTTTAAAA TAAAAATTTT
3700 GCTCATTTAA CGAGTAAATT

3710 AGAAGTTATT TCTTCAATAA
3720 TACTACAAAT ATGATGTTTA
3730 GACTCTACAC CTGAGATGTG
3740 ACACACACAC TGTGTGTGTG
3750 GCGGCGGCGC CGGCGGCGCG
3760 GCGCACACAC TGTGTGTG
3770 ACACACAAAC TGTGTGTTTG
3780 ACACACAAAC TGTGTGTTTG
3790 CTTTTAAAG GAAAAATTTC
3800 AAACGCTAGA TTTGCGATCT

3810 ACCCAACCCC TGGGTTGGGG
3820 CTCTAGGCCA GAGATCCGGT
3830 GAGGAAAACA CTCCTTTGT
3840 TTACAGCTGT AATGTGACA
3850 ATACGCACTT TATGCGTGAA
3860 GTGCCTGTTG CACGGACAAC
3870 CCGTAGAGTA GGCATCTCAT
3880 ATACGGTAGC TATGCCATCG
3890 AGCAGGAGAT TCGTCCTCTA
3900 TACGGTACTA ATGCCATGAT

3910 GCTGGGCTAC CGACCCGATG
3920 TGCCTGAGTT ACGGACTCAA
3930 ACGTCAGCGA TGCAGTCGCT
3940 GAGCTGCAAA CTCGACGTTT
3950 GTTCCTTGCT CAAGGAACGA
3960 ATTCTTTTCT TAAGAAAAGA
3970 GGTGTCGGGG CCACAGCCCC
3980 AGCTGAATAT TCGACTTATA
3990 TAAAAGGGTG ATTTTCCCAC
4000 ATTGTGGAGT TAACACCTCA

4010 TACCGGTTAT ATGGCCAATA
4020 CTGCATTTT GACGTAAAAA
4030 TTTTCTTTTC AAAAGAAAAG
4040 TTATTTTGAC AATAAAACTG
4050 TCTTTTAAA AGAAAAATTT
4060 AAATGCAGGT TTTACGTCCA
4070 AAAGTGACAG TTTCACTGTC
4080 CGGTTCAGGA GCCAAGTCCT
4090 GCTTAAAGAC CGAATTTCTG
4100 ATCAGTGGTG TAGTCACCAC

4110 GAGGGGTGAG CTCCCCACTC
4120 TCAGCGGGTG AGTCGCCCAC
4130 CAAAAGGACA GTTTTCCTGT
4140 AGGATTTGGT TCCTAAACCA
4150 GCCTCGGAGA CGGAGCCTCT
4160 CACGGTCCCC GTGCCAGGGG
4170 TCTCCGCCTC AGAGGCGGAG
4180 CAGAGAAGAG GTCTCTTCTC
4190 CAGGCAGGCA GTCCGTCCGT
4200 GCTCCCGGGA CGAGGGCCCT

4210 CCGAAGCCGG GGCTTCGGCC
4220 GTCCGCATCC CAGGCGTAGG
4230 CCCGGCGGCG GGGCCGCCGC
4240 AGCTGGTGGC TCGACCACCG
4250 TCAGCAGCGG AGTCGTCGCC
4260 CGCTTCAGGT GCGAAGTCCA
4270 GAGTGCGGCG CTCACGGCGG
4280 GGGCCGGCGT CCCGGCCGCA
4290 CCCGCAGGGC GGGCGTCCCG
4300 CGAGTGGGTG GCTCACCCAC

4310 AGGGCAGACC TCCCGTCTGG
4320 TCCCCGCCG AGGGGCGGGC
4330 TCTGGTGAGA AGACCACTCT
4340 CGGAACCCCC GCCTTGGGGG
4350 ACTTTCCCA TGAAAGGGT
4360 GCGCCTCCCG CGCGGAGGGC
4370 CTTTTCCAC GAAAAGGTG
4380 CAGGTTTTAT GTCCAAAATA
4390 ACCGGCCCCT TGGCCGGGGA
4400 CTACCCCACC GATGGGGTGG

4410 CCCGATTCCC GGGCTAAGGG
4420 TTACATCTTC AATGTAGAAG
4430 TGCGAAGTTG ACGCTTCAAC
4440 CCTTCTACTG GGAAGATGAC
4450 TTTTTAACCC AAAAATTGGG
4460 TTTTAACCC AAAAATTGGG
4470 TGTGTTATC ACACAATAG
4480 ACCCTCGAGG TGGGAGCTCC
4490 TAGGAGGAAA ATCCTCCTTT
4500 AGGGTTTCTG TCCCAAAGAC

4510 CAGTGGCACG GTCACCGTGC
4520 TTTTAATAC AAAAATTATG
4530 CACCTGTGAG GTGGACACTC
4540 GTCTCCAACT CAGAGGTTGA
4550 TGCGATTTTA ACGCTAAAAT
4560 AACAAGTGTC TTGTTCACAG
4570 TTGCCCGAGG AACGGGCTCC
4580 TCCCACCTCA AGGGTGGAGT
4590 GGGGCCAACC CCCGGGTTGG
4600 CCAGAAGGCA GGTCTTCCGT

4610 AGGTGGGCAC TCCACCCGTG
4620 TTCCTCACGC AAGGAGTGCG
4630 CGCGCTGTCC GCGCGACAGG
4640 TGCCGAGTCC ACGGCTCAGG
4650 TTCGCAGTTG AAGCGTCAAC
4660 CTGCGGTAGG GACGCCATCC
4670 GGTTTCTTAC CCAAAGAATG
4680 TGGAAACCCA ACCTTTGGGT
4690 GCAGATGGTG CGTCTACCAC
4700 GCCCCAGCC CGGGGGTCGG

4710 CAGAAAATCG GTCTTTAGC
4720 AAGGCGGCCC TTCCGCCGGG
4730 CTGCCCGCTG GACGGGCGAC
4740 GCATGCCGGC CGTACGGCCG
4750 TTAATGTTTA AATTACAAAT
4760 CGCCTGCAAA GCGGACGTTT
4770 ATCCGCAGTG TAGGCGTCAC
4780 ACTGTGCTT TGACAGCGAA
4790 GCAAGCTCC CGTTTCGAGG
4800 CTCTGCGAGAG GAGACGTCTC

4810
4820
4830
4840
4850
4860
4870
4880
4890
4900

Sequence listing (double-stranded DNA, positions 4901–6100). Each position number is followed by its two complementary strands.

(continuation from previous page — partial row, position numbers cut off at top)
GGACGTCCTC / CCTGCAGGAG
CCCCCGCCAG / GGGGGCGGTC
TCCCGCTACG / AGGGCGATGC
GCTGGCAGCT / CGACCGTGGA
GGGTGGCCAA / CCCACCGGTT
CAGTGAGGCT / GTCACTCCGA
GGAGCCCCTC / CCTCGGGGAG
TGGAAAGGCG / ACCTTTCCGC
TCATGGACAG / AGTACCTGTC

4910
ACCTGGGTCG
TGGACCCAGC

4920
CTTTCTGTCT
GAAAGACAGA

4930
TCGGGTCCCT
AGCCCAGGGA

4940
CCCGGCTTCG
GGGCCGAAGC

4950
CTCGGGACCT
GAGCCCTGGA

4960
GGCTCTCAAG
CCGAGAGTTC

4970
CCAGCTTGGC
GGTCGAACCG

4980
TGGTGGACAG
ACCACCTGTC

4990
ACCGGTGCGC
TGGCCACGCG

5000
TCTGCACACC
AGACGTGTGG

5010
CGAGTGCGAA
GCTCACGCTT

5020
TTCCACGGCC
AAGGTGGCCG

5030
GTGAGAATGA
CACTCTTACT

5040
GCGTGCTCGT
CGCACGAGCA

5050
GGTCCTGGCC
CCAGGACCGG

5060
CTGAGGTCCC
GACTCCAGGG

5070
TGGGTCGCAG
ACCCAGCGTC

5080
CTGTTCCCTC
GACAAGGGAG

5090
TCCCAGGCCG
AGGGTCCGGC

5100
CCCCCTCCAG
GGGGGAGGTC

5110
GTGACTGCGA
CACTGACGCT

5120
GGCAACCTGT
CCGTTGGACA

5130
TCTAACGGAA
AGATTGCCTT

5140
ACCGAGTACA
TGGCTCATGT

5150
TCCTCCAGAA
AGGAGGTCTT

5160
TTCCCCGGCT
AAGGGGCCGA

5170
AGGATCCGTG
TCCTAGGCAC

5180
CGACACACTC
GCTGTGTGAG

5190
GCCAGCCGCA
CGGTCGGCGT

5200
GTCGCCCCTC
CAGCGGGGAG

5210
CGGGGCTTCG
GCCCCGAAGC

5220
AGGATTTTAA
TCCTAAAATT

5230
TTTCGTGGTA
AAAGCACCAT

5240
CCTGCGGCTCG
GGACGCCGAGC

5250
AAATCCAGAC
TTTAGGTCTG

5260
TTCGAGCGCT
AAGCTCGCGA

5270
GGAGCCTGGG
CCTCGGACCC

5280
GTTTTGGGGA
CAAAACCCCT

5290
TTTGTTTTTT
AAACAAAAAA

5300
TGTTTGTTTT
ACAAACAAAA

5310
TCGCTTCGGA
AGCGAAGCCT

5320
TCCTGAACTC
AGGACTTGAG

5330
GGGCAGAGGT
CCCGTCTCCA

5340
GACTCAGTAG
CTGAGTCATC

5350
AGTGCGCTAG
TCACGCGATC

5360
GCAGGTTCCC
CGTCCAAGGG

5370
AGTGGTGGGG
TCACCACCCC

5380
GCGGCGAGATG
CGGCGCTCTAC

5390
AGCTCCGAAG
TCGAGGCTTC

5400
TCGCCTCCAC
AGCGGAGGTG

5410
CGCTGCCGGG
GCGACGGCCC

5420
CGAAGCAGCT
GCTTCGTCGA

5430
TCTGGACCGC
AGACCTGGCG

5440
AGAACCAACC
TCTTGGTTGG

5450
CGGCTCCCAA
GCCGAGGGTT

5460
CTGGTGTCCC
GACCACAGGG

5470
CCAACCCGTC
GGTTGGGCAG

5480
AAGCTCAGCA
TTCGAGTCGT

5490
CAGCCTCTTT
GTCGGAGAAA

5500
CCCTGGGGCG
GGGACCCCGC

5510
CCTAGCTCAA
GGATCGAGTT

5520
AGCCGCCTTT
TCGGCGGAAA

5530
CTCTTTGCGC
GAGAAACGCG

5540
TCTTTCAGGT
AGAAAGTCCA

5550
AAACGATGCC
TTTGCTACGG

5560
GGACGCGGTC
CCTGCGCCAG

5570
CCGCAGCCTC
GGCGTCGGAG

5580
CTGGGTCTCA
GACCCAGAGT

5590
GCACATATTC
CGTGTATAAG

5600
CACACCTACG
GTGTGGATGC

5610
TCCCCTGACC
AGGGGACTGG

5620
TGTGCTCCTA
ACACGAGGAT

5630
GAAGCTGGAG
CTTCGACCTC

5640
AGAGCAGGAG
TCTCGTCCTC

5650
CCTTCGGTGG
GGAAGCCACC

5660
GGCAGCTCAA
CCGTCGAGTT

5670
AATGTAGGTA
TTACATCCAT

5680
ACTGCGGGCC
TGACGCCCGG

5690
AGGAGCAGCG
TCCTCGTCGC

5700
CCCAGATGCC
GGGTCTACGG

5710
ATCGGTCCCT
TAGCCAGGGA

5720
GCCTTTGAGC
CGGAAACTCG

5730
GTCGACGGCT
CAGCTGCCGA

5740
GATCTTTTGG
CTAGAAAACC

5750
TTTGAGGGAG
AAACTCCCTC

5760
AGACTGGCGC
TCTGACCGCG

5770
TGGAGTTTTG
ACCTCAAAAC

5780
AATTCCGAAT
TTAAGGCTTA

5790
CATGTGCAGA
GTACACGTCT

5800
ATGTGAATC
TACGACTTAG

5810
TTCCCCCAGC
AAGGGGGTCG

5820
CAGGACGAAT
GTCCTGCTTA

5830
AAGACAGCGC
TTCTGTCGCG

5840
ATTCTCGTAA
TAAGAGCATT

5850
GGAAAAGCAG
CCTTTTCGTC

5860
TTCTGGAATT
AAGACCTTAA

5870
AGGAGTCTCC
TCCTCAGAGG

5880
TGGATCTTCG
ACCTAGAAGC

5890
GCATGTTGCA
CGTACAACGT

5900
CACCCAGCTT
GTGGGTCGAA

5910
CGGGTAGGGA
GCCCATCCCT

5920
GGGAGTCGGG
CCCTCAGCCC

5930
GTCCCGGGCT
CAGGGCCCGA

5940
AGGCCAGCCC
TCCGGTCGGG

5950
GGCAGGTGGA
CCGTCCACCT

5960
GAGGGTCCCC
CTCCCAGGGG

5970
GGCAGCCCCG
CCGTCGGGGC

5980
GGCGGCCCCG
CCGCCGGGGC

5990
GCCATGTCTT
CGGTACAGAA

6000
TAATGCCCTG
ATTACGGGAC

6010
TGGCCTTCTG
CCCCTTCATG

6020
AGGGTTCCCA

6030
GGGCTGGCCA

6040
GGGTTGTTTC

6050
CCACCCGCGC

6060
GCGCGGCTCTC

6070
AAACCCACCT

6080
ACCCCCAGCC

6100
GGCAGGGCTC

Sequence fragments (top row, left to right):

```
GGGGAAGTAC   ACCGGAAGAC   TCCCAAGGGT   CCCGACCGGT   CCCAACAAAG   CGGCGGAGAG   TGGGGGTCGG   TTTGGGTGGA   CCGTCCCGAG
```

```
6110  CCTCCAGCCG   6120  AGACCTTTTG   6130  ATTCCCGGCT   6140  CCCGGCGTCC   6150  CGCCTCCGCG
      GGAGGTCGGC         TCTGGAAAAC         TAAGGGCCGA         GCGGAGGCGG         GCGGAGGCGC

6160  CCAGCCCGGG   6170  AGGTGGCCCT   6180  GGACAGCCGG   6190  ACCTCGCCCG   6200  GCCCCGGCTG
      GGTCGGCGGC         TCCACCGGGA         CCTGTCGGCC         TGGAGCGGGC         CGGGGCCGAC

6210  GGACCATGGT   6220  GTTTCTCTCG   6230  GGAAATGCTT   6240  CCGACAGCTC   6250  CAACTGCACC
      CCTGGTACCA         CAAAGAGAGC         CCTTTACGAA         GGCTGTCGAG         GTTGACGTGG

6260  CAACCGCCGG   6270  CACCGGTGAA   6280  CATTTCCAAG   6290  GCCATTCTGC   6300  TCGGGGTGAT
      GTTGGCGGCC         GTGGCCACTT         GTAAAGGTTC         CGGTAAGACG         AGCCCCACTA

6310  CTTGGGGGGC   6320  CTCATTCTTT   6330  TCGGGGTGCT   6340  GGGTAACATC   6350  TCTCCGTAGC
      GAACCCCCCG         GAGTAAGAAA         AGCCCCACGA         CCCATTGTAG         AGAGGCATCG

6360  CTGTCACCGA   6370  CACCTGCACT   6380  CAGTCACGCA   6390  CTAGTGATCC   6400  CTACTACATC
      GACAGTGGCT         GTGGACGTGA         GTCAGTGCGT         GATCACTAGG         GATGATGTAG

6410  GTCAACCTGG   6420  CGGTGGCCGA   6430  CCTCCTGCTC   6440  ACCTCCACGG   6450  TGCTGCCCTT
      CAGTTGGACC         GCCACCGGCT         GGAGGACGAG         TGGAGGTGCC         ACGACGGGAA

6460  CTCCGCCATC   6470  TTCGAGGTCC   6480  TAGGCTACTG   6490  GGCCTTCGGC   6500  AGGGTCTTCT
      GAGGCGGTAG         AAGCTCCAGG         ATCCGATGAC         CCGGAAGCCG         TCCCAGAAGA

6510  GGCGGCCAGTG  6520  GGCGGCAGTG   6530  GATGTGCTGT   6540  GCTGCACCGC   6550  GTCCATCATG
      CCGCCGGTCAC        CCGCCGTCAC         CTACACGACA         CGACGTGGCG         CAGGTAGTAC

6560  GGCCTCTGCA   6570  TCATCTCCAT   6580  CGACCGCTAC   6590  ATCGGCGTGA   6600  GCTACCCGCT
      CCGGAGACGT         AGTAGAGGTA         GCTGGCGATG         TAGCCGCACT         CGATGGGCGA

6610  GCGCTACCCA   6620  ACCATCGTCA   6630  CCCAGAGGAG   6640  GGGTCTCATG   6650  GCTCTGCTCT
      CGCGGATGGGT        TGGTAGCAGT         GGGTCTCCTC         CCCAGAGTAC         CGAGACGAGA

6660  GCGTCTGGGC   6670  ACTCTCCCTG   6680  GTCATATCCA   6690  TTGGACCCCT   6700  GTTCGGCTGG
      CGCAGACCCG         TGAGAGGGAC         CAGTATAGGT         AACCTGGGGA         CAAGCCGACC

6710  AGGCAGCCGG   6720  CCCCCGAGGA   6730  CGAGACCATC   6740  TGCCAGATCA   6750  ACGAGGAGCC
      TCCGTCGGCC         GGGGGCTCCT         GCTCTGGTAG         ACGGTCTAGT         TGCTCCTCGG

6760  GGGCTACGTG   6770  CTCTTCTCAG   6780  CGCTGGGCTG   6790  CTTCTACCTG   6800  CCTCTGGCCA
      CCCGATGCAC         GAGAAGAGTC         GCGACCCGAC         GAAGATGGAC         GGAGACCGGT

6810  TCATCCTGGT   6820  CATGTACTGC   6830  CGCGTCTACG   6840  TGGTGGCCAA   6850  GAGGGAGAGC
      AGTAGGACCA         GTACATGACG         GCGCAGATGC         ACCACCGGTT         CTCCCTCTCG

6860  CGGGGCCTCA   6870  AGTCTGGCCT   6880  CAAGACCGAC   6890  AAGTCGGACT   6900  CGGAGCAAGT
      GCCCCGGAGT         TCAGACCGGA         GTTCTGGCTG         TTCAGCCTGA         GCCTCGTTCA

6910  GACGCTCCGC   6920  ATCCATCGGA   6930  AAAACGCCCC   6940  GGCAGGAGGC   6950  AGCGGGATGG
      CTGCGAGGCG         TAGGTAGCCT         TTTTGCGGGG         CCGTCCTCCG         TCGCCCTACC

6960  CCAGCGCCAA   6970  GACCAAGACG   6980  CACTTCTCAG   6990  TGAGGCTCCT   7000  CAAGTTCTCC
      GGTCGCGGTT         CTGGTTCTGC         GTGAAGAGTC         ACTCCGAGGA         GTTCAAGAGG

7010  CGGGAGAAGA   7020  AAGCGGCCAA   7030  AACGCTGGGC   7040  ATCGTGGTCG   7050  GCTGCTTCGT
      GCCCTCTTCT         TTCGCCGGTT         TTGCGACCCG         TAGCACCAGC         CGACGAAGCA

7060  CCTCTGCTGG   7070  CTGCCTTTTT   7080  TCTTAGTCAT   7090  GCCCATTGGT   7100  AAGTCTTGAA
      GGAGACGACC         GACGGAAAAA         AGAATCAGTA         CGGGTAACCA         TTCAGAACTT

7110  CACCCCTCAC   7120  TTTAGCATCT   7130  GGGGGTCTTC   7140  ACCCTCCTCG   7150  GCTTCTGTTA
      GTGGGGAGTG         AAATCGTAGA         CCCCCAGAAG         TGGGAGGAGC         CGAAGACAAT

7160  CCCCAGACTC   7170  CCAGTCCGGG   7180  ATGGAAGAGG   7190  AAGGATTAGC   7200  ATTTCAAACA
      GGGGTCTGAG         GGTCAGGCCC         TACCTTCTCC         TTCCTAATCG         TAAAGTTTGT

7210  GCACAGCTCT   7220  AGGGCAATTA   7230  GAAAGGCTC    7240  CCTTGTAGAA   7250  AAGTGAATTT
      CGTGTCGAGA         TCCCGTTAAT         CTTTCCGAG          GGAACATCTT         TTCACTTAAA

7260  TCATTCTCTT   7270  TCTACTCCAG   7280  TCTCATTTAT   7290  ATTAGGTCCT   7300  AGAGCACTTT
      AGTAAGAGAA         AGATGAGGTC         AGAGTAAATA         TAATCCAGGA         TCTCGTGAAA
```

7310 TTCGACTGTA AAGCTGACAT
7320 AAGTGGCTTC TTCACCGAAG
7330 CAACTGATGC GTTGACTACG
7340 AGATTAATTG TCTAATTAAC
7350 GTCTCTTTAA CAGAGAAATT
7360 TAAGAATGTC ATTCTTACAG
7370 AACTTTTCTT TTGAAAAGAA
7380 AATGCCTATA TTACGGATAT
7390 AGCACGTGTT TCGTGCACAA
7400 CAATTTAAAT GTTAAATTTA

7410 GCATCTGCTC CGTAGACGAG
7420 TCTCTAGTCT AGAGATCAGA
7430 CAGAGTCTCC GTCTCAGAGG
7440 ACCAAGTGCT TGGTTCACGA
7450 TAGGCTGACT ATCCGACTGA
7460 GTGGAATGCC CACCTTACGG
7470 ATTTTCACTC TAAAAGTGAG
7480 TGCTACAGAA ACGATGTCTT
7490 TGCAAATTCT ACGTTTAAGA
7500 CTTGGCCTGA GAACCGGACT

7510 AAATAAGTAC TTTATTCATG
7520 CATGCTTATT GTACGAATAA
7530 CTGGACAAAT GACCTGTTTA
7540 GTGTGATTTT CACACTAAAA
7550 ATTATTGCAT TAATAACGTA
7560 TAGGTTATTC ATCCAATAAG
7570 ATAAGGGTTT TATTCCCAAA
7580 GTTATAATGG CAATATTACC
7590 TCTGTTTATG AGACAAATAC
7600 TTCTATATCT AAGATATAGA

7610 GTGCTAANT CACGATTNAA
7620 TATTTCTGG ATAAAGACC
7630 ATTCAGTATG TAAGTCATAC
7640 GAAGGAATTA CTTCCTTAAT
7650 TGGTCAGCCA ACCAGTCGGT
7660 CTNAGAAAAA GANTCTTTTT
7670 AAAATGATTT TTTTACTAAA
7680 TATGTCCAAA ATACAGGTTT
7690 CCAATTTAAG GGTTAAATTC
7700 CCTTAAATAA GGAATTTATT

7710 TTAATCATAG AATTAGTATC
7720 TATTTCCAAT ATAAAGGTTA
7730 AAGTAAATAC TTCATTTATG
7740 TTATTTTTTT AATAAAAAAA
7750 ATTTTAATAA TAAAATTATT
7760 TAAGTATTAA ATTCATAAAT
7770 AAACAAACAC TTTGTTTGTG
7780 TTTCTCTTAT AAAGAGAATA
7790 CCAAAAATCA GGTTTTTAGT
7800 TCCGGGAAAG AGGCCCTTTC

7810 TTACAAGATA AATGTTCTAT
7820 ACACTGTTTG TGTGACAAAC
7830 AAAATTATAC TTTTAATATG
7840 AGTACNCATA TCATGNGTAT
7850 ATGTTACAAA TACAATGTTT
7860 TCCAATTTTT AGGTTAAAAA
7870 GCAAATGCTA CGTTTACGAT
7880 AATTNGCGTT TTAANCGCAA
7890 TGTCAAAATT ACAGTTTTAA
7900 AAATTGCTCA TTTAACGAGT

TC
AG

```
        10        20        30        40        50
ATAAACACTGAGGCTGTGTCTGTTGCATAACACTGCATCAGAGAATAAAA
TATTTGTGACTCCGACACAGACAACGTATTGTGACGTAGTCTCTTATTTT


        60        70        80        90       100
GGCATGTTCAGATAACCGAATTTTAATATGGATTACTTGCATGGATTCCA
CCGTACAAGTCTATTGGCTTAAAATTATACCTAATGAACGTACCTAAGGT


       110       120       130       140       150
ACTTACTTTTCAATTTAGGCAAAACAATTTACATATGTGGACTCAGTCTG
TGAATGAAAAGTTAAATCCGTTTTGTTAAATGTATACACCTGAGTCAGAC


       160       170       180       190       200
AGTTTCACATTTTCATTTGGTAAAACTTCACAGCAGCTGTTGGTCACTGA
TCAAAGTGTAAAAGTAAACCATTTTGAAGTGTCGTCGACAACCAGTGACT


       210       220       230       240       250
GAGCCAGTGCAACCCTACCCACTGGGCCTGCTCCTGTAATTAATGACACA
CTCGGTCACGTTGGGATGGGTGACCCGGACGAGGACATTAATTACTGTGT


       260       270       280       290       300
CGCGGACCAAGTAGGAATGGTCTTTGAAGATATTGCAAAAGGGTGACA
GCGCCTGGTTCATCCTTACCAGAAACTTCTATAACGTTTTCCCACTGT


       310       320       330       340       350
GTCATAGGAGCTAGTCAGTCAAATGTGAGAAACTCATATGTGTTTGGGAT
CAGTATCCTCGATCAGTCAGTTTACACTCTTTGAGTATACACAAACCCTA


       360       370       380       390       400
CATTTTAACCGTTTAAAAATACAGAAAGATGTCTGTTTGATTGTTTTCCT
GTAAAATTGGCAAATTTTTATGTCTTTCTACAGACAAACTAACAAAAGGA


       410       420       430       440       450
AGCCAATTGGCTTGCTGGCTTTCAAATAATATGTATAAATCTGTGTGTTT
TCGGTTAACCGAACGACCGAAAGTTTATTATACATATTTAGACACACAAA


       460       470       480       490       500
TCTTCCAGGGTCTTTCTTCCCTGATTTCAAGCCCTCTGAAACAGTTTTTA
AGAAGGTCCCAGAAAGAAGGGACTAAAGTTCGGGAGACTTTGTCAAAAAT
        G   S   F   F   P   D   F   K   P   S   E   T   V   F>
        ___TRANSLATION OF DEB$HUM A1A-5'+EXON2+_____>


       510       520       530       540       550
AAATAGTATTTTGGCTCGGATATCTAAACAGCTGCATCAACCCCATCATA
TTTATCATAAAACCGAGCCTATAGATTTGTCGACGTAGTTGGGGTAGTAT
K   I   V   F   W   L   G   Y   L   N   S   C   I   N   P   I   I>
    ___TRANSLATION OF DEB$HUM A1A-5'+EXON2+SPLVAR1_____>


       560       570       580       590       600
TACCCATGCTCCAGCCAAGAGTTCAAAAAGGCCTTTCAGAATGTCTTGAG
ATGGGTACGAGGTCGGTTCTCAAGTTTTTCCGGAAAGTCTTACAGAACTC
 Y   P   C   S   S   Q   E   F   K   K   A   F   Q   N   V   L   R>
    ___TRANSLATION OF DEB$HUM A1A-5'+EXON2+SPLVAR1_____>


       610       620       630       640       650
AATCCAGTGTCTCCGCAGAAAGCAGTCTTCCAAACATGCCCTGGGCTACA
TTAGGTCACAGAGGCGTCTTTCGTCAGAAGGTTTGTACGGGACCCGATGT
  I   Q   C   L   R   R   K   Q   S   S   K   H   A   L   G   Y>
    ___TRANSLATION OF DEB$HUM A1A-5'+EXON2+SPLVAR1_____>
```

```
              660        670        680        690        700
CCCTGCACCCGCCCAGCCAGGCCGTGGAAGGGCAACACAAGGACATGGTG
GGGACGTGGGCGGGTCGGTCCGGCACCTTCCCGTTGTGTTCCTGTACCAC
 T   L   H   P   P   S   Q   A   V   E   G   Q   H   K   D   M   V>
___TRANSLATION OF DEB$HUM A1A-5'+EXON2+SPLVAR1____>


              710        720        730        740        750
CGCATCCCCGTGGGATCAAGAGAGACCTTCTACAGGATCTCCAAGACGGA
GCGTAGGGGCACCCTAGTTCTCTCTGGAAGATGTCCTAGAGGTTCTGCCT
   R   I   P   V   G   S   R   E   T   F   Y   R   I   S   K   T   D>
___TRANSLATION OF DEB$HUM A1A-5'+EXON2+SPLVAR1____>


              760        770        780        790        800
TGGCGTTTGTGAATGGAAATTTTTCTCTTCCATGCCCCGTGGATCTGCCA
ACCGCAAACACTTACCTTTAAAAAGAGAAGGTACGGGGCACCTAGACGGT
   G   V   C   E   W   K   F   F   S   S   M   P   R   G   S   A>
___TRANSLATION OF DEB$HUM A1A-5'+EXON2+SPLVAR1____>


              810        820        830        840        850
GGATTACAGTGTCCAAAGACCAATCCTCCTGTACCACAGCCCGGGTGAGA
CCTAATGTCACAGGTTTCTGGTTAGGAGGACATGGTGTCGGGCCCACTCT
 R   I   T   V   S   K   D   Q   S   S   C   T   T   A   R   V   R>
___TRANSLATION OF DEB$HUM A1A-5'+EXON2+SPLVAR1____>


              860        870        880        890        900
AGTAAAAGCTTTTTGCAGGTCTGCTGCTGTGTAGGGCCCTCAACCCCCAG
TCATTTTCGAAAAACGTCCAGACGACGACACATCCCGGGAGTTGGGGGTC
 S   K   S   F   L   Q   V   C   C   C   V   G   P   S   T   P   S>
___TRANSLATION OF DEB$HUM A1A-5'+EXON2+SPLVAR1____>


              910        920        930        940        950
CCTTGACAAGAACCATCAAGTTCCAACCATTAAGGTCCACACCATCTCCC
GGAACTGTTCTTGGTAGTTCAAGGTTGGTAATTCCAGGTGTGGTAGAGGG
   L   D   K   N   H   Q   V   P   T   I   K   V   H   T   I   S>
___TRANSLATION OF DEB$HUM A1A-5'+EXON2+SPLVAR1____>


              960        970        980        990       1000
TCAGTGAGAACGGGGAGGAAGTCTAGGACAGGAAAGATGCAGAGGAAAGG
AGTCACTCTTGCCCCTCCTTCAGATCCTGTCCTTTCTACGTCTCCTTTCC
 L   S   E   N   G   E   E   V   *>
___TRANSLATION OF DEB$____>


             1010       1020       1030       1040       1050
GGAATAATCTTAGGTACCCACCCCACTTCCTTCTCGGAAGGCCAGCTCTT
CCTTATTAGAATCCATGGGTGGGGTGAAGGAAGAGCCTTCCGGTCGAGAA


             1060       1070       1080       1090       1100
CTTGGAGGACAAGACAGGACCAATCAAAGAGGGGACCTGCTGGGAATGGG
GAACCTCCTGTTCTGTCCTGGTTAGTTTCTCCCCTGGACGACCCTTACCC


             1110       1120       1130       1140       1150
GTGGGTGGTAGACCCAACTCATCAGGCAGCGGGTAGGGCACAGGGAAGAG
CACCCACCATCTGGGTTGAGTAGTCCGTCGCCCATCCCGTGTCCCTTCTC


             1160       1170       1180       1190       1200
GGAGGGTGTCTCACAACCAACCAGTTCAGAATGATACGGAACAGCATTTC
CCTCCCACAGAGTGTTGGTTGGTCAAGTCTTACTATGCCTTGTCGTAAAG


             1210       1220       1230       1240       1250
CCTGCAGCTAATGCTTTCTTGGTCACTCTGTGCCCACTTCAACGAAAACC
GGACGTCGATTACGAAAGAACCAGTGAGACACGGGTGAAGTTGCTTTTGG
```

```
        1260      1270      1280      1290      1300
ACCATGGGAAACAGAATTTCATGCACAATCCAAAAGACTATAAATATAGG
TGGTACCCTTTGTCTTAAAGTACGTGTTAGGTTTTCTGATATTTATATCC


        1310      1320      1330      1340      1350
ATTATGATTTCATCATGAATATTTTGAGCACACACTCTAAGTTTGGAGCT
TAATACTAAAGTAGTACTTATAAAACTCGTGTGTGAGATTCAAACCTCGA


        1360      1370      1380      1390      1400
ATTTCTTGATGGAAGTGAGGGGATTTTATTTTCAGGCTGTTCACTTACTG
TAAAGAACTACCTTCACTCCCCTAAAATAAAAGTCCGACAAGTGAATGAC


        1410      1420      1430      1440      1450
CACAGCCATTTCAACATGGCTTACAAAAGCCTTTCTTGACAAATCACTTA
GTGTCGGTAAAGTTGTACCGAATGTTTTCGGAAAGAACTGTTAGTGAAT


        1460      1470      1480      1490      1500
CCTGTTCCAGAACTCTGTTATGAGAATCCAGAGCTTATAATATTTTGNNA
GGACAAGGTCTTGAGACAATACTCTTAGGTCTCGAATATTATAAAACNNT


        1510      1520      1530      1540      1550
GGCAAAAGATTGTCTCCCATTACTTCTTATCTGCTTCACTATTGCATAAT
CCGTTTTCTAACAGAGGGTAATGAAGAATAGACGAAGTGATAACGTATTA


        1560      1570      1580      1590      1600
GAATGAGCTTCACCTGTGGCATGTTGGAATGAGCCTTATGATCCAAGTAC
CTTACTCGAAGTGGACACCGTACAACCTTACTCGGAATACTAGGTTCATG


        1610      1620      1630      1640      1650
ATTATTCCCGAACTTTGNAAAATACTAATGCTTAGCTTCAGACAATACTG
TAATAAGGGCTTGAAACNTTTTATGATTACGAATCGAAGTCTGTTATGAC


        1660      1670      1680      1690      1700
ATGGNTCNCCAAAGNACTGTCTAACGCAGGAGTTTNCNAACATTTTTTGA
TACCNAGNGGTTTCNTGACAGATTGCGTCCTCAAANGNTTGTAAAAAACT


        1710      1720      1730      1740      1750
TAGGAGGCCATTTGTTCTGGTAAAAGATCCAGTAGTCAACTCAACTTCAT
ATCCTCCGGTAAACAAGACCATTTTCTAGGTCATCAGTTGAGTTGAAGTA


        1760      1770      1780      1790      1800
CTATTCCTACTTTTCTGCAAGAGCTTGGGGNACATGCTATATTTTGCTTT
GATAAGGATGAAAAGACGTTCTCGAACCCCNTGTACGATATAAAACGAAA


        1810      1820      1830      1840      1850
ATGTATACCNAATTTTGGTAAACCATAATAACTCAGTAAAAAAAAAGTCG
TACATATGGNTTAAAACCATTTGGTATTATTGAGTCATTTTTTTTTCAGC


        1860      1870      1880      1890      1900
ACGCGGCCGCGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG
TGCGCCGGCGCTTAAGCTATAGTTCGAATAGCTATGGCAGCTGGAGCTCC


        1910
GGGGGCCCGGTACCCAA
CCCCCGGGCCATGGGTT
```

```
          10        20        30        40        50
AAGACNAGTTGTGTTTTGGGTAAATNAAANANTCTCNTNNNGATTTTTGN
TTCTGNTCAACACAAAACCCATTTANTTTNTNAGAGNANNNCTAAAAACN


          60        70        80        90       100
TNAGGCCCTACAGTNTGCCAGGNATCTTCCAGGANTTTTGAACCCATTGC
ANTCCGGGATGTCANACGGTCCNTAGAAGGTCCTNAAAACTTGGGTAACG


         110       120       130       140       150
CTCTAAANTCNTTGGAAACTCTTGTACCCCTATTTTAACAGANAAAAAAT
GAGATTTNAGNAACCTTTGAGAACATGGGGATAAAATTGTCTNTTTTTA


         160       170       180       190       200
CTGATCATTGTANGGAGNTTAAGGACTTGCCCAAGGCCAAGGGGACCCAT
GACTAGTAACATNCCTCNAATTCCTGAACGGGTTCCGGTTCCCCTGGGTA


         210       220       230       240       250
GATTTAAACCTGGTCTCCCTATTCTCAACNTGCACATTTTCCATAGCCCC
CTAAATTTGGACCAGAGGGATAAGAGTTGNACGTGTAAAAGGTATCGGGG


         260       270       280       290       300
CCCTTCCTCAAGAGAATGGGGGTAAACGTTTTCCCATTTGGATTAGGTGT
GGGAAGGAGTTCTCTTACCCCCATTTGCAAAAGGGTAAACCTAATCCACA


         310       320       330       340       350
GCTGAGGATGCCAAGCCATATCCAAACTTTTTAATGTTCTGTTTCCTTGA
CGACTCCTACGGTTCGGTATAGGTTTGAAAAATTACAAGACAAAGGAACT


         360       370       380       390       400
GATTTGCCTCCAAATTAACCATGGCAAACATGCTGCCAAATCTCCAGCCC
CTAAACGGAGGTTTAATTGGTACCGTTTGTACGACGGTTTAGAGGTCGGG


         410       420       430       440       450
AGTCAAACCCAAGAGCAGGGCATCTGTGGAAGAATTTGGTGTCTGCACTT
TCAGTTTGGGTTCTCGTCCCGTAGACACCTTCTTAAACCACAGACGTGAA


         460       470        480       490       500
CTATCAGACTGCATCACATTTTTAGGAGTTAATGGCCTGGAATGTGATTA
GATAGTCTGACGTAGTGTAAAAATCCTCAATTACCGGACCTTACACTAAT


         510       520       530       540       550
AGGCCTTGCAAGGAGGACATACTATGGCACGCTGGGGGAAGATGGGCATA
TCCGGAACGTTCCTCCTGTATGATACCGTGCGACCCCCTTCTACCCGTAT


         560       570       580       590       600
GAAGAGTATGCAGAAGGGGCCACATTGGCCAAGAACAGTAAAATGCAGTT
CTTCTCATACGTCTTCCCCGGTGTAACCGGTTCTTGTCATTTTACGTCAA


         610       620       630       640       650
GCTGACAGGACACATATCGGGTGTTGTATTGAAGTTATTGATGACCAACC
CGACTGTCCTGTGTATAGCCCACAACATAACTTCAATAACTACTGGTTGG


         660       670       680       690       700
ACAGTTCATAGAAACACTTTTGGGAAGTACATCCCTTTTAAAAATAAATG
TGTCAAGTATCTTTGTGAAAACCCTTCATGTAGGGAAAATTTTTATTTAC


         710       720       730       740       750
AAAGCAAATACGGCATAACTCACTCTCACTCACCTGTATTCCAACTTTTT
TTTCGTTTATGCCGTATTGAGTGAGAGTGAGTGGACATAAGGTTGAAAAA
```

```
        760       770       780       790        800
TTTTGTTTGGACAGAGGGGAATGGATTGTAGATATTTCACCAAGAATTGC
AAAACAAACCTGTCTCCCCTTACCTAACATCTATAAAGTGGTTCTTAACG
                R   G   M   D   C   R   Y   F   T   K   N   C>
                ___TRANSLATION OF DEB$HUM A1A-5'____>


        810       820       830       840       850
AGAGAGCATATCAAGCATGTGAATTTTATGATGCCACCGTGGAGAAAGGG
TCTCTCGTATAGTTCGTACACTTAAAATACTACGGTGGCACCTCTTTCCC
   R   E   H   I   K   H   V   N   F   M   M   P   P   W   R   K   G>
   ___TRANSLATION OF DEB$HUM A1A-5'+SPLICE VAR4+.____>


        860       870       880       890       900
TTCAGAATGCTGATCTCCAGGTAGCTGGAGACCTAGGCAGTCTGCAAATG
AAGTCTTACGACTAGAGGTCCATCGACCTCTGGATCCGTCAGACGTTTAC
    S   E   C   *>
    _____>


        910       920       930       940       950
AGGAGTCAGCTGGAAGCTATGGCTATGTATTATGTGACATCGCTTGTTCC
TCCTCAGTCGACCTTCGATACCGATACATAATACACTGTAGCGAACAAGG


        960       970       980       990      1000
TAATTTCTTTTCACACAAGTGAAAACTGGATATCCCAACCTTCTGGCCCA
ATTAAAGAAAAGTGTGTTCACTTTTGACCTATAGGGTTGGAAGACCGGGT


       1010      1020      1030      1040      1050
GTAGGTTTCATGGTTAAGACCTGGTAGTGAGAACATTTTAGGAACTATTT
CATCCAAAGTACCAATTCTGGACCATCACTCTTGTAAAATCCTTGATAAA


       1060      1070      1080      1090      1100
GCTTGGGCAGGTAATTTTTCACTCTGATCACAGCTACTTAATCAGAGCTT
CGAACCCGTCCATTAAAAAGTGAGACTAGTGTCGATGAATTAGTCTCGAA


       1110      1120      1130      1140      1150
GACAAACTTTCTCAATTGCTTCTGGGGCTGGTCTGCTCAGGTCCTTTGGC
CTGTTTGAAAGAGTTAACGAAGACCCCGACCAGACGAGTCCAGGAAACCG


       1160      1170      1180      1190      1200
CAAAAAAGATGCTGCCTCTGTGTGTGAATACTTGTTGACTTAATTAAAGA
GTTTTTTCTACGACGGAGACACACACTTATGAACAACTGAATTAATTTCT


       1210      1220      1230      1240      1250
AAGAGCTCTGCTCATTAGCAAAGGGCACTGNNGCAGATGGGAGGTAAACT
TTCTCGAGACGAGTAATCGTTTCCCGTGACNNCGTCTACCCTCCATTTGA


       1260      1270      1280      1290      1300
CTCCAGGGAAAAACCAAGTGAAAAGAAAGCAGAGGAGGCAAATATGGAGA
GAGGTCCCTTTTTGGTTCACTTTTCTTTCGTCTCCTCCGTTTATACCTCT


       1310      1320      1330      1340      1350
CATCAGAGGTATGCCTACCAGTTACTCTGATTTTTTTACACTACTAGGAC
GTAGTCTCCATACGGATGGTCAATGAGACTAAAAAAATGTGATGATCCTG


       1360      1370      1380      1390      1400
TTTTAACTATGAAACCACCTGCGCACAGCTCCAGTGGAGCCCAGTTGGAA
AAAATTGATACTTTGGTGGACGCGTGTCGAGGTCACCTCGGGTCAACCTT


       1410      1420      1430      1440      1450
CACTGTTTGCACCTGCCCCGTATCTGCAGACTGTCCTGGGGAGCTGGGCT
GTGACAAACGTGGACGGGGCATAGACGTCTGACAGGACCCCTCGACCCGA
```

```
         1460      1470      1480      1490      1500
GAGCCAGGCTGCTGTGGTGCCATCCATCTTCAGAAAGCAATGGCAGCTGT
CTCGGTCCGACGACACCACGGTAGGTAGAAGTCTTTCGTTACCGTCGACA


         1510      1520      1530      1540      1550
GGCCCTCCTGGCCTCCANAACCCCTGGGGAGCAAGNATGAGTGGGAGATN
CCGGGAGGACCGGAGGTNTTGGGGACCCCTCGTTCNTACTCACCCTCTAN


         1560      1570      1580      1590      1600
ATCACTGGTGGGGGCTNAGGCTGACNAAAAGCAAGTTTAGGAATTTCAAT
TAGTGACCACCCCCGANTCCGACTGNTTTTCGTTCAAATCCTTAAAGTTA


         1610      1620      1630      1640      1650
NGGGGGGGGACAACAAGCCCTNGCNCCCATCNAGCAATTAGGTCAATTCAC
NCCCCCCCCTGTTGTTCGGGANCGNGGGTAGNTCGTTAATCCAGTTAAGTG


         1660      1670      1680      1690      1700
GCCNCCCCAAGACCCAAATNTGGGNGGAGGGGGGTTNNAGANTTTGGCCCT
CGGNGGGGGTTCTGGGTTTANACCCNCCTCCCCCAANNTCTNAAACCGGGA


TC
AG
```

```
          10        20        30        40        50
CCCTCATTGCNAACANTTGAAGCATNTTNAAGGNGTGGCNTNCTTTTTGT
GGGAGTAACGNTTGTNAACTTCGTANAANTTCCNCACCGNANGAAAAACA


          60        70        80        90       100
AACCAAAACCTGATTGGAAGCAGAAGTTGGGGGGGGGGGGAAATTGGAAG
TTGGTTTTGGACTAACCTTCGTCTTCAACCCCCCCCCCCCTTTAACCTTC


         110       120       130       140       150
CAAAAGGGACCGAACTTGAAGCNTGTACTNCCCAGACTTCTCATTGGAAG
GTTTTCCCTGGCTTGAACTTCGNACATGANGGGTCTGAAGAGTAACCTTC


         160       170       180       190       200
CTCCAGGTCACAATCTTAGTNCTAATTTCAGGTTCCTGCCCAGTCGAGTC
GAGGTCCAGTGTTAGAATCANGATTAAAGTCCAAGGACGGGTCAGCTCAG


         210       220       230       240       250
TAGACATTNCTGGGGCACCCTTTAAGTGGTCTCCAGCACCCTTCTAGGCA
ATCTGTAANGACCCCGTGGGAAATTCACCAGAGGTCGTGGGAAGATCCGT


         260       270       280       290       300
GGTGGCTTCACAGGGAAGTCCACACTGCAAGCTCAGCTCACCACGACTCA
CCACCGAAGTGTCCCTTCAGGTGTGACGTTCGAGTCGAGTGGTGCTGAGT


         310       320       330       340       350
GGCTGACGGGGTAGTCAGCCATGCTCGGAGGCTGAATTTGGCAGGGACTT
CCGACTGCCCCATCAGTCGGTACGAGCCTCCGACTTAAACCGTCCCTGAA


         360       370       380       390       400
GCTGCCATCTCCCTGCCAAATGGTCTCTCGTGACTCAGAATCTCAAACTT
CGACGGTAGAGGGACGGTTTACCAGAGAGCACTGAGTCTTAGAGTTTGAA


         410       420       430       440       450
GTTTTAAAGAGAGGAAAAAAGTCACTTTCGGGGATGAGGTTCTTGGCCCA
CAAAATTTCTCTCCTTTTTTCAGTGAAAGCCCCTACTCCAAGAACCGGGT


         460       470       480       490       500
ACTCTGCTTTATATAAACACAGTCTATGGCTATTTCAGTCTTCTGGATTT
TGAGACGAAATATATTTGTGTCAGATACCGATAAAGTCAGAAGACCTAAA


         510       520       530       540       550
TGAGAAGCAGCTGCAAGGATGAACGGATTGGTGTTGGCCCAAATTAAAAA
ACTCTTCGTCGACGTTCCTACTTGCCTAACCACAACCGGGTTTAATTTTT


         560       570       580       590       600
GAAGAGTATTCAGTTCTTTCAGTGTTTGGAGAAAGAAGACCAAAAGCATC
CTTCTCATAAGTCAAGAAAGTCACAAACCTCTTTCTTCTGGTTTTCGTAG


         610       620       630       640       650
ATCTCACAGGGAGCAGAATGTGACCAGCCTGGCTAATGAGGAAATGAGAG
TAGAGTGTCCCTCGTCTTACACTGGTCGGACCGATTACTCCTTTACTCTC


         660       670       680       690       700
GGATCCTCAACTTGAGAACCCGCTCTACTGAAGTCTGAACTTGGAAAAAT
```

CCTAGGAGTTGAACTCTTGGGCGAGATGACTTCAGACTTGAACCTTTTTA

```
         710       720       730       740       750
GGACACATTGGGTTTGGAGTAAGAATTCTTACTCTACAAAAGGATAAAAT
CCTGTGTAACCCAAACCTCATTCTTAAGAATGAGATGTTTTCCTATTTTA

         760       770       780       790       800
TGTGATCACATTGATGCATGATGCCTAGGATATTAAAAATGCATGATTAA
ACACTAGTGTAACTACGTACTACGGATCCTATAATTTTTACGTACTAATT

         810       820       830       840       850
TTAAATGTTAGTCTACCTTGTGTTTTAAAGGGACACACACCCATGACATG
AATTTACAATCAGATGGAACACAAAATTTCCCTGTGTGTGGGTACTGTAC
                                G   H   T   P   M   T>
                                ___TRANSLATION____>

         860       870       880       890       900
AAGCCAGCTTCCCGTCCACGACTGTTGTCCTTACTGCCCAAGGAAGGGGA
TTCGGTCGAAGGGCAGGTGCTGACAACAGGAATGACGGGTTCCTTCCCCT

         910       920       930       940       950
GCATGAAACCCACCACTGGTCCTGCGACCCACTGTCTTTGGAATCCACCC
CGTACTTTGGGTGGTGACCAGGACGCTGGGTGACAGAAACCTTAGGTGGG

         960       970       980       990       1000
CAGGAGCCCAGGAGCCTTGCCTGACACTTGGATTTACTTCTTTATCAAGC
GTCCTCGGGTCCTCGGAACGGACTGTGAACCTAAATGAAGAAATAGTTCG

        1010      1020      1030      1040      1050
ATCCATCTGACTAAGGCACAAATCCAACATGTTACTGTTACTGATACAGG
TAGGTAGACTGATTCCGTGTTTAGGTTGTACAATGACAATGACTATGTCC

        1060      1070      1080      1090      1100
AAAAACAGTAACTTAAGGAATGATCATGAATGCAAAGGGAAAGAGGAAAA
TTTTTGTCATTGAATTCCTTACTAGTACTTACGTTTCCCTTTCTCCTTTT

        1110      1120      1130      1140      1150
GAGCCTTCAGGGACAAATAGCTCGATTTTTTGTAAATCAGTTTCATACAA
CTCGGAAGTCCCTGTTTATCGAGCTAAAAAACATTTAGTCAAAGTATGTT

        1160      1170      1180      1190      1200
CCTCCCTCCCCCATTTCATTCTTAAAAGTTAATTGAGAATCATCAGCCAC
GGAGGGAGGGGGTAAAGTAAGAATTTTCAATTAACTCTTAGTAGTCGGTG

        1210
GTGTAGGGTGTGAG
CACATCCCACACTC
```

Table 6. Localization of naturally occurring single nucleotide polymorphisms (SNPs) in human $\alpha_{1a}AR$ coding region

Amino Acid Position	Amino Acid Change	SNP Name (dbSNP ID, handle, release date)	Nucleotide Position	Nucleotide Change	Domain Position	Allelic f(-)				
						All	All (Except Coriell)	Black	Hispanic	Caucasian
5	None		15	G→C	N-terminus	0.002	0.003	0.000	0.000	0.005
154	Ser→Ala	S154A (G2286a1*)	460	T→G	TM4	0.007	0.005	0.000	0.013	0.005
166	Arg→Lys	R166K	497	G→A	TM4	0.002	0.003	0.000	0.000	0.005
200	Ile→Ser	I200S (rs2229125, WICVAR, 8/15/01)	599	T→G	TM5	0.005	0.005	0.000	0.000	0.010
247	Gly→Arg	G247R (rs3730287, WIPGA, 8/28/02)	739	G→A	IL3	0.003	0.003	0.000	0.023	0.000
311	Val→Ile	V311I	931	G→A	TM7	0.002	0.000	0.000	0.000	0.000
347	Arg→Cys	R347C (rs1048101, TSC-CSHL, 1/29/01)	1039	C→T	C-terminus	0.463	0.482	0.267	0.525	0.565
401	None		1203	T→G	C-terminus	0.002	0.003	0.012	0.000	0.000
465	Glu→Asp	E465D (rs2229126, WICVAR, 8/15/2001)	1395	A→T	C-terminus	0.030	0.021	0.023	0.025	0.020

Nucleotide number is based on the first nucleotide of the start codon being +1 of human $\alpha_{1a}AR$ cDNA (GeneBank accession number is L31774).

TM: transmembrane region; IL: intracellular loop. SNPs are identified from 281 individuals: Coriell n=90 (enriched for minorities but exact race

not known), Black n=43, Hispanic n=40, caucasian n=101. *It is a Whitehead Internal ID cited from website http//www.cardiogenomics.org.

TABLE 7. Agonist and Antagonist Binding Affinities of Human α_{1a}AR Wild Type and SNPs from High-expression Stable Clones

	α_{1a}AR WT	S154A	R166K	I200S	G247R	V311I	R347C	E465D
Domain location	—	TM4	TM4	TM5	IL3	TM7	C-terminus	C-terminus
Expression (pmol/mg)	1.77±0.24	1.53±0.07	2.37±0.28	1.53±0.10	1.96±0.40	1.75±0.37	1.76±0.22	2.02±0.27
[^{125}I]HEAT(K_d, pmol/L)	42.1±6.5	38.1±4.4	35.9±10.0	34.5±1.3	32.1±5.3	42.7±16.5	39.9±3.2	38.8±3.7
Agonists (pK_i)								
Norepinephrine	4.73±0.03	4.73±0.08	4.34±0.10**	4.90±0.08	4.61±0.03	4.34±0.06**	4.73±0.03	4.64±0.02
Epinephrine	5.01±0.02	4.87±0.06	4.64±0.04**	5.18±0.08	4.96±0.01	4.57±0.08**	5.01±0.05	4.91±0.04
Phenylephrine	4.69±0.04	4.60±0.08	4.44±0.01*	4.82±0.07	4.58±0.05	4.27±0.02**	4.69±0.04	4.75±0.06
Oxymetazoline	7.67±0.02	7.57±0.02	7.61±0.04	7.84±0.10	7.64±0.03	7.82±0.03	7.58±0.07	7.57±0.08
Antagonists (pK_i)								
Prazosin	9.39±0.10	9.34±0.09	9.19±0.08	9.69±0.10	9.40±0.14	9.20±0.06	9.30±0.07	9.61±0.10
Phentolamine	7.70±0.03	7.77±0.05	7.77±0.04	7.20±0.05**	7.60±0.07	7.60±0.04	7.75±0.01	7.84±0.01
5-Methylurapidil	8.44±0.04	8.45±0.10	8.38±0.08	8.40±0.07	8.45±0.07	8.87±0.06**	8.39±0.02	8.44±0.04

pK_i values for the binding of agonists and antagonists were determined in competition binding experiments on rat-1 cell membranes stably transfected to express α_{1a}AR WT or its SNPs (receptor density > 1.5 pmol/mg protein). K_d values for the antagonist [^{125}I]HEAT and receptor densities were determined from saturation binding studies. *$P<0.05$, **$P<0.01$ compared with α_{1a}AR WT. Data are reported as the mean±SEM of 3–7 different experiments.

TABLE Σ. Agonist and Antagonist Binding Affinities of Human α_{1a}AR Wild Type and SNPs from Low-expression Stable Clones

	α_{1a}AR WT	S154A	R166K	I200S	G247R	V311I	R347C	E465D
Domain location	—	TM4	TM4	TM5	IL3	TM7	C-terminus	C-terminus
Expression (pmol/mg)	0.36±0.01	0.37±0.09	0.44±0.05	0.21±0.06	0.33±0.03	0.29±0.03	0.36±0.01	0.26±0.05
[^{125}I]HEAT(K_d, pmol/L)	42.9±4.0	37.7±16.3	42.6±6.1	49.9±7.7	49.1±10.8	49.7±10.0	50.7±9.8	46.6±16.4
Agonists (pK_i)								
Norepinephrine	4.73±0.08	4.78±0.02	4.31±0.04**	4.82±0.07	4.77±0.09	4.32±0.04**	4.65±0.04	4.65±0.04
Epinephrine	5.02±0.07	4.97±0.02	4.62±0.03**	5.15±0.03	5.09±0.13	4.36±0.01**	4.97±0.08	4.90±0.10
Phenylephrine	4.68±0.08	4.72±0.07	4.37±0.04**	4.75±0.03	4.60±0.08	4.24±0.03**	4.62±0.03	4.66±0.11
Oxymetazoline	7.65±0.02	7.78±0.03	7.63±0.06	7.65±0.01	7.56±0.04	7.78±0.02	7.75±0.02	7.76±0.09
Antagonists (pK_i)								
Prazosin	9.51±0.11	9.36±0.09	9.32±0.08	9.54±0.04	9.40±0.14	9.50±0.03	9.47±0.12	9.22±0.11
Phentolamine	7.68±0.03	7.64±0.01	7.73±0.04	7.31±0.07**	7.60±0.07	7.49±0.10	7.70±0.05	7.76±0.11
5-Methylurapidil	8.42±0.11	8.38±0.07	8.41±0.15	8.43±0.07	8.45±0.07	8.94±0.06**	8.56±0.04	8.48±0.04

pK_i values for the binding of agonists and antagonists were determined in competition binding experiments on rat-1 cell membranes stably transfected to express α_{1a}AR WT or its SNPs (receptor density < 0.5 pmol/mg protein). K_d values for the antagonist [^{125}I]HEAT and receptor densities were determined from saturation binding studies. **$P<0.01$ compared with α_{1a}AR WT. Data are reported as the mean±SEM of 3-5 different experiments.

Table 9 Functional alterations and clinical implications of human α_{1a}AR SNPs

SNP	Functional Alteration(s)	Possible Biological Mechanisms	Clinical Implications
R166K	Decreases binding affinity for endogenous agonists, modestly reduces potency of NE in stimulating IP$_3$ formation	Hampers agonist binding to residue F163	1) The homozygous form might provide a novel mechanism underlying human hypotension syndromes, 2) may protect against sympathetically-mediated hypertension
I200S	Decreases binding affinity for antagonist phentolamine	Influences three consecutive residues (Q177, I178, N179) involved in phentolamine binding	May explain variable individual response to drug treatments
G247E	Enhances receptor activity and stimulates cell growth	Unknown	1) Might be relevant to cardiovascular remodeling (e.g. in hypertension, atherosclerosis, myocardial hypertrophy) and benign prostatic hyperplasia, 2) May be important in neoplastic transformation (e.g. some cancers such as prostate cancer)
V311I	Decreases binding affinity for endogenous agonists, reduces potency of NE in stimulating IP$_3$ formation	Stabilizes the salt-bridge	Same as SNP R166K